

Random walks colliding before getting trapped

Louigi Addario-Berry* Roberto I. Oliveira† Yuval Peres‡ Perla Sousi§

Abstract

Let P be the transition matrix of a finite, irreducible and reversible Markov chain. We say the continuous time Markov chain X has transition matrix P and speed λ if it jumps at rate λ according to the matrix P . Fix $\lambda_X, \lambda_Y, \lambda_Z \geq 0$, then let X, Y and Z be independent Markov chains with transition matrix P and speeds λ_X, λ_Y and λ_Z respectively, all started from the stationary distribution. What is the chance that X and Y meet before either of them collides with Z ? For each choice of λ_X, λ_Y and λ_Z with $\max(\lambda_X, \lambda_Y) > 0$, we prove a lower bound for this probability which is uniform over all transitive, irreducible and reversible chains. In the case that $\lambda_X = \lambda_Y = 1$ and $\lambda_Z = 0$ we prove a strengthening of our main theorem using a martingale argument. We provide an example showing the transitivity assumption cannot be removed for general λ_X, λ_Y and λ_Z .

Keywords and phrases. Transitive Markov chains, mixing time, hitting times, martingale.
MSC 2010 subject classifications. Primary 60J10.

1 Introduction

Consider three independent random walks X, Y, Z over the same finite connected graph. What is the probability that X, Y meet at the same vertex before either of them meets Z ? If the initial distributions of the three walkers are the same, this probability is at least $1/3$ by symmetry, at least if we assume that ties (i.e. triple meetings) are broken symmetrically.

Now consider a similar problem where the initial states X_0, Y_0, Z_0 are all sampled independently from the same distribution, but Z stays put while X and Y move. What is the probability that X and Y meet before hitting Z ?

There are several examples of bounds [1, 4, 5] relating the meeting time of two random walks to the hitting time of a fixed vertex by a single random walk. These typically provide upper bounds for meeting times in terms of worst-case or average hitting times, sometimes up to constant factors. In light of this, it seems natural to conjecture that the probability in question is at least $1/3$. However, the previous argument by symmetry fails. In fact, to the best of our knowledge, no known universal lower bound for this probability is known.

It will be convenient to consider the problem in continuous time. For the remainder of the paper let P be the transition matrix of an irreducible and reversible Markov chain on a finite state space with stationary distribution π . Let X and Y be two independent continuous time Markov chains

*McGill University, Montréal, Canada; louigi.addario@mcgill.ca

†IMPA, Rio de Janeiro, Brazil; rimfo@impa.br. Supported by a *Bolsa de Produtividade em Pesquisa* from CNPq.

‡Microsoft Research, Redmond, Washington, USA; peres@microsoft.com

§University of Cambridge, Cambridge, UK; p.sousi@statslab.cam.ac.uk

that jump at rate 1 according to the transition matrix P and let $Z \sim \pi$ be independent of X and Y .

We define $M^{X,Y}$ to be the first time X and Y meet, i.e.

$$M^{X,Y} = \inf\{t \geq 0 : X_t = Y_t\}.$$

We also define:

$$M^{W,Z} = \inf\{t \geq 0 : X_t = Z_t\} \quad (W \in \{X, Y\}).$$

We write $M^{\text{good}} = M^{X,Y}$ and $M^{\text{bad}} = M^{X,Z} \wedge M^{Y,Z}$.

1.1 Main results

Our first result proves a universal lower bound on the probability $\mathbb{P}(M^{\text{good}} < M^{\text{bad}})$ for the class of transitive chains. First we recall the definition.

Definition 1.1. Fix a chain with transition matrix P and state space E . An *automorphism* of P is a bijection $\varphi : E \rightarrow E$ such that $P(z, w) = P(\varphi(z), \varphi(w))$ for all z, w . The chain P is *transitive* if for all $x, y \in E$ there exists an automorphism φ of P with $\varphi(x) = y$.

Theorem 1.2. *Let P be the transition matrix of a finite irreducible and reversible chain with two or more states. Assume X_0 and Y_0 are independent with law π . If P is transitive, then*

$$\mathbb{P}(M^{\text{good}} < M^{\text{bad}}) \geq \frac{1}{4}.$$

Next we consider a more general setup. We say that a random walk W has speed λ_W and transition matrix P , if it jumps at rate λ_W according to the matrix P .

Suppose again that P is the transition matrix of an irreducible and reversible Markov chain on a finite state space with stationary distribution π . Let $\lambda_X = 1, 0 \leq \lambda_Y \leq 1$ and $0 \leq \lambda_Z < \infty$. Let X, Y and Z be three independent continuous time Markov chains with speeds λ_X, λ_Y and λ_Z respectively and transition matrix P .

For the remainder of the paper, we write \mathbb{P} for the probability measure under which X_0, Y_0 and Z_0 are independent with law π . We also write $\mathbb{P}_{a,b,c}$ in the case when $(X_0, Y_0, Z_0) = (a, b, c)$. For computations that only involve two chains we drop one index writing only $\mathbb{P}_{a,b}$; which two chains are involved will always be clear from context. Likewise, we write \mathbb{P}_a when only one chain is involved. We define $M^{X,Y}$ as above and redefine:

$$M^{W,Z} = \inf\{t \geq 0 : X_t = Z_t\} \quad (W \in \{X, Y\}).$$

$M^{\text{good}} = M^{X,Y}$ and $M^{\text{bad}} = M^{X,Z} \wedge M^{Y,Z}$ are defined as before. Again we are interested in uniform lower bounds on the probability of the event $\{M^{\text{good}} < M^{\text{bad}}\}$ that have good dependence on the three speeds.

Theorem 1.3. *There exists $c > 0$ such that the following holds. Let P be the transition matrix of a transitive and reversible chain with stationary distribution π and at least two states. Suppose that X, Y and Z are three independent continuous time Markov chains with speeds $\lambda_X = 1, \lambda_Y \leq 1$ and $0 \leq \lambda_Z < \infty$ and transition matrix P started from π . Then*

$$\mathbb{P}(M^{\text{good}} < M^{\text{bad}}) \geq \frac{c}{(\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2}.$$

The proof shows that we may take $c = 1/4752$, which implies a version of Theorem 1.2 with $1/4$ replaced by $1/((\sqrt{2} + 1)^2 \cdot 4752)$. The constant c most likely can be improved, but the dependence of the lower bound on λ_Z is sharp when $\lambda_Z \nearrow +\infty$. Indeed, if P is simple random walk over a large complete graph with n vertices, then

$$\mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) = \frac{1 + \lambda_Y}{2(1 + \lambda_Y + \lambda_Z)} - O\left(\frac{1}{n}\right),$$

where the term $O(1/n)$ corresponds to the possibility of meetings at time 0.

It is natural to ask if our theorems can be extended to all transitive chains. The next theorem shows that the answer is no for the more general Theorem 1.3. The theorem essentially asserts that there are graphs where typical meeting times are much smaller than typical hitting times.

Theorem 1.4. *For all $\varepsilon > 0$ there exists a finite connected graph G such that if P corresponds to simple random walk on G and $\lambda_X = 1, \lambda_Y = 0$ and $\lambda_Z = 1$, then $\mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) < \varepsilon$.*

On the other hand, we believe that for certain values of λ_X, λ_Y and λ_Z , universal lower bounds are possible without transitivity. Here is a concrete conjecture, which relates to the setting of Theorem 1.2.

Conjecture 1.5. If $\lambda_Y = \lambda_X = 1$ and $\lambda_Z = 0$, the inequality

$$\mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) \geq 1/3$$

holds for all finite irreducible and reversible chains P .

Alexander Holroyd (personal communication) pointed out an example showing that for any $\delta > 0$ there exist transitive chains for which $\mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) \leq 1/3 + \delta$. We describe this example in Section 6. This means that, if true, Conjecture 1.5 is best possible even for transitive chains. However, we note that any uniform lower bound

$$\mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) \geq c > 0$$

for all P , and for λ_X, λ_Y and λ_Z as in Conjecture 1.5, would be a new result.

Remark 1.6. Without reversibility, the conjecture fails badly. Consider a clockwise continuous time random walk on a cycle of length $2n$. More precisely, with $P = (p_{ij})_{1 \leq i, j \leq n}$ we have $p_{ij} = 1$ if $j = (i + 1) \bmod n$ and $p_{ij} = 0$ otherwise. The distance between independent random walkers behaves as continuous time simple symmetric random walk reflected at 0 and n . So started from stationarity, it typically takes such walkers time of order n^2 to meet. On the other hand, the hitting time of any point is at most of order n .

Before we continue, we say a few words about the main proof ideas. The unifying theme of the proofs of Theorems 1.2 and 1.3 is the relationship between meeting times and hitting times of single vertices when P is transitive. Aldous and Fill [1, Chapter 14/Proposition 5 and Chapter 3] have related the expected values of these random variables via martingales, and we use similar ideas to prove Theorem 1.2.

For the proof of Theorem 1.3, we need a stronger result establishing *identities in distribution* of meeting and hitting times, which (somewhat surprisingly) seems to be new: see Lemma 3.5 below. The proof of Theorem 1.3 requires several other tools, including small time estimates for hitting times (Lemma 3.4) and an occupation time formula for product chains (Lemma 4.2).

The proof of Theorem 1.4 builds a graph with two parts: the “Up” part concentrates the bulk of the stationary measure, but the “Down” part is where meetings tend to happen, and they happen fast. As a result, only a negligible fraction of the “Up” part is explored before X and Z meet, and the upshot is that $M^{X,Y} > M^{X,Z}$ with high probability.

2 The 1/4 lower bound

In this section we prove Theorem 1.2. The argument is fairly short, and much simpler than the one for the more general Theorem 1.3. Before presenting the proof, we recall some standard facts about hitting times which are also used later on.

The hitting time of a state $z \in \Omega$ by X is the first time t at which $X_t = z$, i.e.

$$\tau_z^X := \inf\{t \geq 0 : X_t = z\}. \quad (2.1)$$

We define τ_z^Y similarly and we also let

$$t_{\text{hit}}^* := \max_{x \in \Omega} \mathbb{E}_x[\tau_x^X] \quad \text{and} \quad t_{\text{hit}} := \max_{(x,z) \in \Omega} \mathbb{E}_x[\tau_z^X]. \quad (2.2)$$

Whenever there is no confusion, i.e. if there is a single chain in question, we will drop the dependence on X or Y from the notation of the hitting times.

Lemma 2.1. *For any reversible chain with two or more states we have*

$$0 < t_{\text{hit}} \leq 2t_{\text{hit}}^*.$$

Moreover, if X is a reversible and transitive chain, then for all $x, z \in \Omega$ and all $t \geq 0$ we have

$$\mathbb{P}_x(\tau_z \leq t) = \mathbb{P}_z(\tau_x \leq t).$$

Proof. For a proof of the first assertion in discrete time we refer the reader to [3, Lemma 10.2]. A proof of the symmetry property specific to transitive chains can be found in [1, Lemma 1, Chapter 7]. \square

Lemma 2.2 (Aldous). *Let P be an irreducible and reversible transition matrix. Suppose that X and Y are independent continuous time Markov chains that jump at rate 1 according to the transition matrix P . For all $x, z \in \Omega$ define $f(x, z) := \mathbb{E}_x[\tau_z^X]$. Then $f(X_t, z) + t$, $f(Y_t, z) + t$ and $f(X_t, Y_t) + 2t$ are martingales up to time*

$$S := M^{\text{good}} \wedge M^{\text{bad}} = M^{X,Y} \wedge \tau_z^X \wedge \tau_z^Y,$$

for any initial states $(x, y) \in \Omega^2$ and any $z \in \Omega$.

For a proof of the lemma above we refer the reader to [1, Chapter 14/Proposition 5 and Chapter 3].

Proof of Theorem 1.2. Since X and Y are two independent copies of the same chain, we have $\mathbb{E}_a[\tau_b^X] = \mathbb{E}_a[\tau_b^Y]$ for all a, b . By Lemma 2.2 we now get that $(G_t)_{t \geq 0}$ is a martingale up to time S , where

$$G_t := \mathbb{E}_{X_t}[\tau_z^X] + \mathbb{E}_{Y_t}[\tau_z^X] - \mathbb{E}_{X_t}[\tau_{Y_t}^X] \quad (t \geq 0).$$

This martingale is bounded (because the state space is finite). The fact that the chain is finite and irreducible implies $S < \infty$ almost surely for all initial states. We deduce from optional stopping that

$$\mathbb{E}[G_0] = \mathbb{E}[G_S].$$

The left hand side above is given by the quantity t_{hit}^* defined in (2.2). This is because

$$\mathbb{E}[G_0] = \mathbb{E}_\pi[\tau_z^X] + \mathbb{E}_\pi[\tau_z^X] - \sum_{y \in \Omega} \pi(y) \mathbb{E}_\pi[\tau_y^X] = \mathbb{E}_\pi[\tau_z^X], \quad (2.3)$$

where the second equality follows from the fact that for a transitive chain, $\mathbb{E}_\pi[\tau_y^X]$ is independent of y . Using this a second time yields

$$t_{\text{hit}}^* = \mathbb{E}[G_S]. \quad (2.4)$$

On the other hand, at time S we have two alternatives.

- If $\tau_z^X \wedge \tau_z^Y \leq M^{\text{good}}$, either $X_S = z$, and then $G_S = \mathbb{E}_{Y_S}[\tau_z^X] - \mathbb{E}_z[\tau_{Y_S}^X]$, or $Y_S = z$, in which case $G_S = \mathbb{E}_{X_S}[\tau_z^X] - \mathbb{E}_{X_S}[\tau_z^X]$. In both cases $G_S = 0$: this is obviously true in the second case, and follows from Lemma 2.1 in the first case.
- On the other hand, if $M^{\text{good}} < \tau_z^X \wedge \tau_z^Y$, then $G_S = 2\mathbb{E}_{X_S}[\tau_z^X] \leq 2t_{\text{hit}}$.

We deduce that

$$G_S \leq 2t_{\text{hit}} \mathbf{1}(M^{\text{good}} < \tau_z^X \wedge \tau_z^Y).$$

Plugging this into (2.4) gives

$$t_{\text{hit}}^* \leq 2t_{\text{hit}} \mathbb{P}(M^{\text{good}} < M^{\text{bad}}).$$

Using that $t_{\text{hit}} \leq 2t_{\text{hit}}^*$ from Lemma 2.1 finishes the proof. \square

3 Towards the general lower bound

In this section we collect the tools that we will use in the proof of Theorem 1.3. We first argue that the obvious “fix” to the proof of Theorem 1.2 does not work in all cases. Indeed, a straightforward extension of Lemma 2.2 establishes that

$$G_t := f(X_t, Z_t) + f(Y_t, Z_t) - \frac{1 + \lambda_Y + 2\lambda_Z}{1 + \lambda_Y} f(X_t, Y_t)$$

is a martingale up to time S . One can see that in this case

$$\mathbb{E}[G_0] = \left(1 - \frac{2\lambda_Z}{1 + \lambda_Y}\right) t_{\text{hit}}^*,$$

which easily yields

$$\mathbb{P}(M^{\text{good}} < M^{\text{bad}}) \geq \frac{1}{4} \cdot \left(1 - \frac{2\lambda_Z}{1 + \lambda_Y}\right).$$

In particular, we obtain the same bound as in Theorem 1.2 provided that $\lambda_Z = 0$ and $\lambda_X \neq 0$. However, this bound becomes useless when $\lambda_Z > (1 + \lambda_Y)/2$. Other linear combinations of $f(X_t, Z_t)$, $f(Y_t, Z_t)$ and $f(X_t, Y_t)$ also fail to achieve our goal when λ_Z is large. So, in general a different strategy is needed.

3.1 Hitting times for states and trajectories

In this section we collect some results on hitting times for a single Markov chain.

Recall the quantity t_{hit}^* defined in (2.2). The next lemma shows that, up to a constant factor, t_{hit}^* also bounds expected hitting times of moving targets, from arbitrary initial states. This lemma is essentially due to Oliveira [4, Lemma 1.1], but in this particular form it appeared in [5].

Lemma 3.1. *Let X be a reversible Markov chain taking values in Ω and $\mathbf{h} = (h_t)_{t \geq 0}$ a deterministic, càdlàg, Ω -valued trajectory. If*

$$\tau_{\mathbf{h}} := \inf\{t \geq 0 : X_t = h_t\},$$

then for any $x \in \Omega$,

$$\mathbb{E}_x[\tau_{\mathbf{h}}] \leq 11 t_{\text{hit}}^*.$$

Proof. In [5] using [4, Lemma 1.1] it is proved that

$$\mathbb{E}_x[\tau_{\mathbf{h}}] \leq c t_{\text{hit}}.$$

for a universal constant $c > 0$, where t_{hit} is as in (2.2). Inspection of the proof [5] shows that $c \leq 4 + 5/4$, therefore $2c \leq 11$. Lemma 2.1 finishes the proof. \square

For a reversible transition matrix P we let $\lambda_* = \max_{i \geq 2} |\lambda_i|$, where $1 = \lambda_1 \geq \lambda_2 \geq \lambda_3, \dots$ are its eigenvalues in decreasing order. We define the relaxation time via $t_{\text{rel}} = (1 - \lambda_*)^{-1}$.

The next lemma bounds the probability that τ_z is small for reversible chains. Clearly, if no assumption is made on z and on the starting state x , such an estimate cannot be very good (think of two adjacent points on a path). The next lemma shows that if we choose the “right” starting state, and only consider the majority of possible target states, we can show that τ_z dominates an exponential random variable with mean t_{rel} . This will be used later to upper bound the probability that τ_z/t_{hit} is very small.

Lemma 3.2. *Let X be a reversible chain. There exist $x \in \Omega$ and a subset $A \subset \Omega$ with stationary measure $\pi(A) \geq 1/2$ such that, if $\tau_A := \min_{z \in A} \tau_z$, then for any $t > 0$,*

$$\mathbb{P}_x(\tau_A > t) \geq e^{-\frac{t}{t_{\text{rel}}}}.$$

Proof. We use the spectral theory of reversible chains [1, Section 3.4]. The first step is to note that P has a non-zero eigenfunction $\varphi : \Omega \rightarrow \mathbb{R}$ such that

$$P\varphi = \left(1 - \frac{1}{t_{\text{rel}}}\right) \varphi.$$

This eigenfunction is orthogonal to the constant eigenfunction in the inner product induced by π , so it must take both positive and negative values. We may assume without loss of generality that the set

$$A := \{z \in \Omega : \varphi(z) \leq 0\}$$

has measure $\pi(A) \geq 1/2$ (if that is not the case, replace φ with $-\varphi$). Choose $x \in \Omega$ with $\varphi(x) > 0$ as large as possible. On one hand

$$\forall t > 0, u \in \Omega : \mathbb{E}_u[\varphi(X_t)] = [e^{t(P-I)} \varphi](u) = e^{-\frac{t}{t_{\text{rel}}}} \varphi(u). \quad (3.1)$$

In particular, for all $u \in A$ we have $\mathbb{E}_u[\varphi(X_t)] \leq 0$. Since $X_{\tau_A} \in A$, the strong Markov property then gives

$$\begin{aligned}\mathbb{E}_x[\varphi(X_t) \mathbf{1}(\tau_A \leq t)] &= \mathbb{E}_x \left[\mathbb{E}_{X_{\tau_A}}[\varphi(X_{t-\tau_A})] \mathbf{1}(\tau_A \leq t) \right] \\ &\leq \mathbb{E}_x[0 \cdot \mathbf{1}(\tau_A \leq t)] = 0.\end{aligned}$$

Plugging this into (3.1) with the choice $u = x$, and recalling $\varphi(X_t) \leq \varphi(x)$ always, we obtain

$$e^{-\frac{t}{t_{\text{rel}}}} \varphi(x) = \mathbb{E}_x[\varphi(X_t)] \leq \mathbb{E}_x[\varphi(X_t) \mathbf{1}(\tau_A > t)] \leq \varphi(x) \mathbb{P}_x(\tau_A > t).$$

Dividing both sides by $\varphi(x)$ (which is > 0) finishes the proof. \square

3.2 Results for transitive chains

In this section we prove results on hitting times under the assumption that P is transitive and reversible.

3.2.1 A small-time estimate for hitting times

We start by recalling a result which follows from the complete positivity of the law of τ_z when starting from π together with bounds from Aldous and Brown [2].

Lemma 3.3 (Aldous-Brown). *Let P be an irreducible and reversible chain. Define $f(s) = \mathbb{E}_\pi[\tau_z - s \mid \tau_z > s]$ for all $s > 0$. Then f is an increasing function and*

$$\sup_s f(s) \leq \mathbb{E}_\pi[\tau_z] + t_{\text{rel}}.$$

Proof. Let $Q = I - P$ and Q_z be the restriction of Q to $\Omega \setminus \{z\}$. We recall the complete positivity of the law of τ_z starting from π (see for instance [2, eqn. (18)]): there exist non-negative constants $(p_i, 1 \leq i \leq n)$ such that for all t

$$\mathbb{P}_\pi(\tau_z > t) = \sum_{i=1}^m p_i e^{-\gamma_i t},$$

where $0 < \gamma_1 < \dots < \gamma_m$ are the distinct eigenvalues of $-Q_z$. Note that $\gamma_1^{-1} = \mathbb{E}_\alpha[\tau_z]$, where α is any quasistationary distribution on $\Omega \setminus \{z\}$ corresponding to the eigenvalue γ_1 .

Using the above representation we can rewrite f as follows

$$f(s) = \int_s^\infty \frac{\mathbb{P}_\pi(\tau_z > t)}{\mathbb{P}_\pi(\tau_z > s)} dt = \frac{\sum_{i=1}^m p_i e^{-\gamma_i s} / \gamma_i}{\sum_{i=1}^m p_i e^{-\gamma_i s}}.$$

A straightforward differentiation now gives that f is increasing. From the above expression we also deduce

$$\lim_{s \rightarrow \infty} f(s) = \frac{1}{\gamma_1} = \mathbb{E}_\alpha[\tau_z]. \quad (3.2)$$

From [2, Corollary 4] we have

$$\mathbb{E}_\alpha[\tau_z] \leq \mathbb{E}_\pi[\tau_z] + t_{\text{rel}}.$$

Therefore, using this and the fact that f is increasing and (3.2) we conclude that for all s

$$f(s) \leq \mathbb{E}_\alpha[\tau_z] \leq \mathbb{E}_\pi[\tau_z] + t_{\text{rel}}$$

which completes the proof. \square

The next lemma essentially improves upon Lemma 3.2 from Section 3.1.

Lemma 3.4. *Suppose that P is reversible and transitive. Then for any $x \in \Omega$, there exists a subset $A_x \subset \Omega$ with $\pi(A_x) \geq 1/2$ such that, for any $\theta > 0$,*

$$\frac{1}{\pi(A_x)} \sum_{z \in A_x} \pi(z) \mathbb{P}_x(\tau_z \leq \theta t_{\text{hit}}) \leq 6\sqrt{\theta}$$

Proof. By Lemma 3.2 there exists $x \in \Omega$ and a set $A = A_x$ with $\pi(A_x) \geq 1/2$ such that

$$\mathbb{P}_x(\tau_A > t) \geq e^{-\frac{t}{t_{\text{rel}}}}.$$

Since the chain is transitive, this in fact holds for all x . We now fix $x \in \Omega$ and $\theta > 0$.

We will consider two cases separately: $t_{\text{rel}} < \sqrt{\theta} t_{\text{hit}}$ and $t_{\text{rel}} \geq \sqrt{\theta} t_{\text{hit}}$.

Suppose first that $t_{\text{rel}} \geq \sqrt{\theta} t_{\text{hit}}$. Since $\mathbb{P}_x(\tau_z > t) \geq \mathbb{P}_x(\tau_A > t)$ for all $z \in A$, we obtain

$$\frac{1}{\pi(A)} \sum_{z \in A} \pi(z) \mathbb{P}_x(\tau_z > \theta t_{\text{hit}}) \geq e^{-\frac{\theta t_{\text{hit}}}{t_{\text{rel}}}} \geq 1 - \frac{\theta t_{\text{hit}}}{t_{\text{rel}}} \geq 1 - \sqrt{\theta}$$

which concludes the proof in this case.

Suppose next that $t_{\text{rel}} < \sqrt{\theta} t_{\text{hit}}$. In this case it suffices to prove

$$\forall z \in \Omega : \mathbb{P}_\pi(\tau_z > \theta t_{\text{hit}}) \geq 1 - 3\sqrt{\theta}. \quad (3.3)$$

To see that this suffices, we use the fact that P is transitive and apply Lemma 2.1 to obtain that $\mathbb{P}_x(\tau_z > t)$ is symmetric in x and z . As a result, (3.3) implies

$$1 - 3\sqrt{\theta} \leq \mathbb{P}_\pi(\tau_x > \theta t_{\text{hit}}) = \sum_{z \in \Omega} \pi(z) \mathbb{P}_z(\tau_x > \theta t_{\text{hit}}) = \sum_{z \in \Omega} \pi(z) \mathbb{P}_x(\tau_z > \theta t_{\text{hit}}).$$

Therefore we obtain

$$\sum_z \pi(z) \mathbb{P}_x(\tau_z \leq \theta t_{\text{hit}}) \leq 3\sqrt{\theta}.$$

Since $\pi(A) \geq 1/2$, we conclude

$$\frac{1}{\pi(A)} \sum_{z \in A} \pi(z) \mathbb{P}_x(\tau_z \leq \theta t_{\text{hit}}) \leq 6\sqrt{\theta}.$$

It remains to prove (3.3). Since P is transitive, $\mathbb{E}_\pi[\tau_z] = t_{\text{hit}}^*$ is independent of z . Moreover, we are assuming that $t_{\text{rel}} \leq \sqrt{\theta} t_{\text{hit}}$, and hence using Lemma 2.1 we get that $t_{\text{rel}} \leq 2\sqrt{\theta} t_{\text{hit}}^*$ so

$$\mathbb{E}_\pi[\tau_z] + t_{\text{rel}} \leq (1 + 2\sqrt{\theta}) t_{\text{hit}}^*. \quad (3.4)$$

We now obtain for all $s \geq 0$ and for all $z \in \Omega$

$$\begin{aligned} t_{\text{hit}}^* = \mathbb{E}_\pi[\tau_z] &\leq s + \mathbb{E}_\pi[\tau_z - s \mid \tau_z > s] \mathbb{P}_\pi(\tau_z > s) \leq s + (\mathbb{E}_\pi[\tau_z] + t_{\text{rel}}) \mathbb{P}_\pi(\tau_z > s) \\ &\leq s + (1 + 2\sqrt{\theta}) t_{\text{hit}}^* \mathbb{P}_\pi(\tau_z > s), \end{aligned}$$

where the second inequality follows from Lemma 3.3 and the final one from (3.4). Taking $s = \theta t_{\text{hit}}^*$ gives

$$\mathbb{P}_\pi(\tau_z > \theta t_{\text{hit}}^*) \geq \frac{1 - \theta}{1 + 2\sqrt{\theta}} \geq 1 - 2\sqrt{\theta}.$$

This now finishes the proof of (3.3), since for all $z \in \Omega$ we have

$$\mathbb{P}_\pi(\tau_z > \theta t_{\text{hit}}) \geq \mathbb{P}_\pi(\tau_z > 2\theta t_{\text{hit}}^*) \geq 1 - 2\sqrt{2}\sqrt{\theta} \geq 1 - 3\sqrt{\theta},$$

where for the first inequality we used again Lemma 2.1. □

3.2.2 Distributional identities for meeting and hitting times

Our next result shows that the distributions of hitting and meeting times are intimately related for transitive and reversible P .

Lemma 3.5. *Let P be a reversible and transitive transition matrix. Let X, Y and Z be three independent continuous time Markov chains with speeds $\lambda_X = 1, \lambda_Y \geq 0$ and $\lambda_Z \geq 0$ and transition matrix P . Then for all $(x, z) \in \Omega^2$, the distribution of $\frac{\tau_z^X}{\lambda_Y + \lambda_Z}$ under \mathbb{P}_x is the same as the distribution of $M^{Y,Z}$ under $\mathbb{P}_{(x,z)}$.*

A special case of this lemma is when $\lambda_X = \lambda_Z = 1$, in which case we obtain the following corollary of Lemma 3.5.

$$\mathbb{E}_{(x,z)} [M^{X,Z}] = \frac{1}{2} \mathbb{E}_x [\tau_z^X].$$

This equality is well known and is usually proven by martingale methods such as the ones used in the proof of Theorem 1.2. Somewhat oddly, it seems that Lemma 3.5 is new, or at least was not widely known before. We also note the following corollary of Lemma 3.4 and Lemma 3.5.

Corollary 3.6. *Let P be a transitive and reversible transition matrix and let X, Y and Z be three independent continuous time Markov chains with speeds $\lambda_X = 1$ and $\lambda_Y, \lambda_Z \geq 0$ respectively and transition matrix P . Then for all $x \in \Omega$ there exists a subset $A_x \subset \Omega$ with $\pi(A_x) \geq 1/2$ such that:*

$$\frac{1}{\pi(A_x)} \sum_{z \in A_x} \pi(z) \mathbb{P}_{(x,z)} (M^{X,Z} \leq \theta t_{\text{hit}}) \leq 6\sqrt{(1 + \lambda_Z)} \theta$$

and

$$\frac{1}{\pi(A_x)} \sum_{z \in A_x} \pi(z) \mathbb{P}_{(x,z)} (M^{Y,Z} \leq \theta t_{\text{hit}}) \leq 6\sqrt{(\lambda_Y + \lambda_Z)} \theta.$$

Proof of Lemma 3.5. Define the functions

$$g_{(x,z)}(t) := \mathbb{P}_x (\tau_z \leq (\lambda_Y + \lambda_Z) t) \text{ and } f_{x,z}(t) := \mathbb{P}_{(x,z)} (M^{Y,Z} \leq t) \text{ } ((x, z) \in \Omega^2, t \geq 0).$$

We will be done once we show that $g_{x,z}(t) = f_{x,z}(t)$ for all $(x, z) \in \Omega^2$ and $t \geq 0$. These equalities are true (by inspection) when $t = 0$. We are going to show that the functions $(f_{(x,z)}(\cdot))_{(x,z) \in \Omega^2}$ and $(g_{x,z}(\cdot))_{(x,z) \in \Omega^2}$ satisfy the same linear system of ordinary differential equations (with the derivatives at $t = 0$ interpreted as right derivatives). Then the equality for all $t \geq 0$ will follow from the general uniqueness theory of linear ODE's.

To prove that f and g satisfy the same system of ODE's, we will use a standard formula for the cumulative distribution function of a hitting time. If $(V_t)_{t \geq 0}$ is an irreducible continuous time Markov chain over a set Ω_V with transition rates $q(v, w)$, and $A \subset \Omega_V$ is a nonempty subset of the state space, the hitting time τ_A^V of A by V satisfies

$$\frac{d}{dt} \mathbb{P}_v (\tau_A^V \leq t) = \begin{cases} 0, & v \in A; \\ \sum_{w \in V} q(v, w) (\mathbb{P}_w (\tau_A^V \leq t) - \mathbb{P}_v (\tau_A^V \leq t)), & v \in \Omega_V \setminus A. \end{cases} \quad (3.5)$$

The derivative is understood as a right derivative at time $t = 0$.

Let $\Delta = \{(x, x) : x \in \Omega\} \subset \Omega^2$ be the diagonal set. We first apply (3.5) to the product chain $(V_t)_{t \geq 0} = (Y_t, Z_t)_{t \geq 0}$, with $\Omega_V = \Omega^2$, and $A = \Delta$. In this case $\tau_A^V = M^{Y,Z}$, and a straightforward computation with the transition rates gives:

$$\frac{d}{dt} f_{x,z}(t) = \begin{cases} 0, & x = z \\ \lambda_Y \sum_{x' \in \Omega} P(x, x') (f_{(x',z)}(t) - f_{(x,z)}(t)) \\ + \lambda_Z \sum_{z' \in \Omega} P(z, z') (f_{(x,z')} (t) - f_{(x,z)}(t)), & x \neq z. \end{cases} \quad (3.6)$$

We now apply the same formula (3.5) with $(V_t)_{t \geq 0} = (X_t)_{t \geq 0}$. Note that $g_{(x,z)}(t) := \mathbb{P}_x(\tau_z \leq s(t))$ where $s(t) = (\lambda_Y + \lambda_Z)t$, so the chain rule implies

$$\frac{d}{dt} g_{(x,z)}(t) = \begin{cases} 0, & x = z \\ (\lambda_Y + \lambda_Z) \sum_{x' \in \Omega} P(x, x') (g_{(x',z)}(t) - g_{(x,z)}(t)), & x \neq z. \end{cases} \quad (3.7)$$

We will now make crucial use of transitivity, which allows us to use Lemma 2.1 to deduce that $\mathbb{P}_x(\tau_z^X \leq (\lambda_Y + \lambda_Z)t)$ is symmetric in x and z , i.e.

$$\mathbb{P}_x(\tau_z^X \leq (\lambda_Y + \lambda_Z)t) = \mathbb{P}_z(\tau_x^X \leq (\lambda_Y + \lambda_Z)t),$$

that is $g_{x,z}(\cdot) = g_{z,x}(\cdot)$ for all x, z . This allows us to reverse the roles of x and z in (3.7) to obtain:

$$\frac{d}{dt} g_{(x,z)}(t) = \begin{cases} 0, & x = z \\ (\lambda_Y + \lambda_Z) \sum_{z' \in \Omega} P(z, z') (g_{(x,z')}(t) - g_{(x,z)}(t)), & x \neq z. \end{cases} \quad (3.8)$$

We add the two formulas (3.7) and (3.8) with weights $\lambda_Y/(\lambda_Y + \lambda_Z)$ and $\lambda_Z/(\lambda_Y + \lambda_Z)$ respectively. The upshot is:

$$\frac{d}{dt} g_{(x,z)}(t) = \begin{cases} 0, & x = z \\ \lambda_Y \sum_{x' \in \Omega} P(x, x') (g_{(x',z)}(t) - g_{(x,z)}(t)) \\ + \lambda_Z \sum_{z' \in \Omega} P(z, z') (g_{(x,z')}(t) - g_{(x,z)}(t)), & x \neq z. \end{cases}$$

This is precisely the system of ODEs we obtained for the f 's in (3.6) and it concludes the proof. \square

4 The general lower bound

In this section we prove Theorem 1.3.

Proof of Theorem 1.3. We let $n = |\Omega|$ denote the number of states. The transitivity assumption implies $\pi(v) = 1/n$ for all $v \in \Omega$. The next lemma will be used in the proof. We defer its proof until Section 4.1.

Lemma 4.1. *Let P be a reversible and transitive transition matrix. Let X, Y and Z be three independent continuous time chains with transition matrix P and speeds $\lambda_X = 1$ and $\lambda_Y, \lambda_Z \geq 0$. Let μ be the probability measure given by*

$$\mu(A) := \mathbb{P}\left((X_{M^{\text{good}}}, Y_{M^{\text{good}}}, Z_{M^{\text{good}}}) \in A \mid M^{\text{good}} < M^{\text{bad}}\right) \quad \text{for } A \subseteq \Omega \times \Omega \times \Omega. \quad (4.1)$$

Then

$$\int_0^{+\infty} \mathbb{P}_\mu(X_t = Y_t, t < M^{\text{bad}}) dt \leq \frac{22 t_{\text{hit}}}{n}.$$

Our proof is based on the analysis of the time that X, Y spend on the diagonal $\Delta = \{(x, x) : x \in \Omega\}$ prior to time M^{bad} , i.e.

$$T := \int_0^{M^{\text{bad}}} \mathbf{1}(X_t = Y_t) dt = \int_0^\infty \mathbf{1}(X_t = Y_t, t < M^{\text{bad}}) dt. \quad (4.2)$$

Note that $M^{\text{good}} < M^{\text{bad}}$ if and only if $T > 0$, so

$$\mathbb{E}[T] = \mathbb{E}\left[T \mid M^{\text{good}} < M^{\text{bad}}\right] \mathbb{P}(M^{\text{good}} < M^{\text{bad}})$$

$$\begin{aligned}
&= \mathbb{E}_\mu \left[\int_0^\infty \mathbf{1}(X_t = Y_t, t < M^{\text{bad}}) dt \right] \mathbb{P}(M^{\text{good}} < M^{\text{bad}}) \\
&= \mathbb{P}(M^{\text{good}} < M^{\text{bad}}) \int_0^\infty \mathbb{P}_\mu(X_t = Y_t, t < M^{\text{bad}}) dt.
\end{aligned} \tag{4.3}$$

Lemma 4.1 upper bounds the integral appearing above. Thus in order to obtain a lower bound for $\mathbb{P}(M^{\text{good}} < M^{\text{bad}})$ it suffices to lower bound

$$\mathbb{E}[T] = \int_0^\infty \mathbb{P}(X_t = Y_t, t < M^{\text{bad}}) dt. \tag{4.4}$$

Using reversibility and the fact that π is uniform for a transitive chain we obtain

$$\begin{aligned}
\mathbb{P}(X_t = Y_t, t < M^{\text{bad}}) &= \mathbb{P}(X_t = Y_t, \forall s \leq t, X_s \neq Z_s \text{ and } Y_s \neq Z_s) \\
&= \mathbb{P}(X_0 = Y_0, \forall s \leq t, X_s \neq Z_s \text{ and } Y_s \neq Z_s) \\
&= \sum_{x \in \Omega, z \in \Omega \setminus \{x\}} \frac{\mathbb{P}_{(x,x,z)}(\forall s \leq t, X_s \neq Z_s \text{ and } Y_s \neq Z_s)}{n^3} \\
&= \sum_{x \in \Omega, z \in \Omega \setminus \{x\}} \frac{\mathbb{P}_{(x,x,z)}(M^{\text{bad}} > t)}{n^3} = \frac{1}{n} \sum_{x,z \in \Omega} \pi(x) \pi(z) \mathbb{P}_{(x,x,z)}(M^{\text{bad}} > t).
\end{aligned}$$

Plugging this back into (4.4) gives

$$\mathbb{E}[T] = \frac{1}{n} \sum_{x \in \Omega} \pi(x) \left(\int_0^\infty \sum_{z \in \Omega} \pi(z) \mathbb{P}_{(x,x,z)}(M^{\text{bad}} > t) dt \right). \tag{4.5}$$

At this point we recall that $M^{\text{bad}} = M^{X,Y} \wedge M^{Y,Z}$, therefore

$$\forall (x, z) \in \Omega^2; \mathbb{P}_{(x,x,z)}(M^{\text{bad}} > t) \geq 1 - \mathbb{P}_{(x,z)}(M^{X,Z} \leq t) - \mathbb{P}_{(x,z)}(M^{Y,Z} \leq t).$$

By Corollary 3.6, for each $x \in \Omega$ there exists a subset A_x with $\pi(A_x) \geq 1/2$ for which we have the following bound for all $t > 0$

$$\sum_{z \in A_x} \pi(z) \mathbb{P}_{(x,x,z)}(M^{\text{bad}} > t) \geq \pi(A_x) \left[1 - 6(\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z}) \sqrt{\frac{t}{t_{\text{hit}}}} \right].$$

Therefore, fixing $\xi > 0$ and $x \in \Omega$:

$$\begin{aligned}
\int_0^\infty \sum_{z \in \Omega} \pi(z) \mathbb{P}_{(x,x,z)}(M^{\text{bad}} > t) dt &\geq \frac{1}{2} \int_0^\xi \sum_{z \in A_x} \frac{\pi(z)}{\pi(A_x)} \mathbb{P}_{(x,x,z)}(M^{\text{bad}} > t) dt \\
&\geq \frac{1}{2} \int_0^\xi \left[1 - 6(\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z}) \sqrt{\frac{t}{t_{\text{hit}}}} \right] dt \\
&= \frac{\xi}{2} \left[1 - 4(\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z}) \sqrt{\frac{\xi}{t_{\text{hit}}}} \right].
\end{aligned}$$

We can maximize the right hand side by taking

$$\xi := \frac{t_{\text{hit}}}{36(\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2},$$

which gives the bound

$$\forall x \in \Omega : \int_0^\infty \sum_{z \in \Omega} \pi(z) \mathbb{P}_{(x,x,z)} \left(M^{\text{bad}} > t \right) dt \geq \frac{t_{\text{hit}}}{216 (\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2}.$$

Combining this with (4.5) gives that

$$\mathbb{E}[T] \geq \frac{t_{\text{hit}}}{216 n (\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2}.$$

Therefore, using this together with Lemma 4.1 and (4.3) we conclude

$$\mathbb{P} \left(M^{\text{good}} < M^{\text{bad}} \right) \geq \frac{1}{4752 (\sqrt{1 + \lambda_Z} + \sqrt{\lambda_Y + \lambda_Z})^2}$$

and this finishes the proof. \square

4.1 Occupation time up to a stopping time

The goal of this section is to prove Lemma 4.1. We start with a more general setting. We give the proof of the lemma at the end of the section.

It is well known that a finite irreducible chain $(V_t)_{t \geq 0}$ with state space Ω_V , started from a point x and stopped at a stopping time $\tau > 0$ with $V_\tau = x$, satisfies:

$$\forall v \in \Omega_V : \mathbb{E}_x \left[\int_0^\tau \mathbf{1}(X_t = v) dt \right] = \pi_V(v) \mathbb{E}_x [\tau],$$

where π_V is the unique stationary measure of $(V_t)_{t \geq 0}$ (some simple conditions on τ are necessary for this). There are also extensions of this lemma to the case where V_0 and V_τ are not necessarily equal, but have the same distribution [1, Proposition 2.4, Chapter 2]. The following lemma extends this idea even further, and shows that τ may be a stopping time for a “larger” Markov chain.

Lemma 4.2. *Suppose $(U_t)_{t \geq 0}$, $(V_t)_{t \geq 0}$ are independent, irreducible, continuous time Markov chains with finite state spaces Ω_U and Ω_V respectively. Assume μ is a probability measure over $\Omega_U \times \Omega_V$ and that τ is a stopping time for the process $(U_t, V_t)_{t \geq 0}$ with the following properties.*

- (1) $\mathbb{P}_\mu(\tau > 0) = 1$;
- (2) $\mathbb{E}_\mu[\tau] < \infty$.
- (3) $\mathbb{P}_\mu(V_0 = \cdot) = \mathbb{P}_\mu(V_\tau = \cdot)$.

Then for all $v \in \Omega_V$

$$\mathbb{E}_\mu \left[\int_0^\tau \mathbf{1}(V_t = v) dt \right] = \mathbb{E}_\mu[\tau] \pi_V(v)$$

where π_V is the stationary distribution of V .

Proof. In this proof we will interchange integrals, expectations and summations several times. Instead of justifying this at each step, we note right away that all of these interchanges are valid, because the integrands are non-negative.

Consider the row vector h with nonnegative coordinates

$$h(v) := \mathbb{E}_\mu \left[\int_0^\tau \mathbf{1}(V_t = v) dt \right] = \int_0^\infty \mathbb{P}_\mu(V_t = v, \tau > t) dt \quad (v \in \Omega_V).$$

Note that $\sum_{v \in V} h(v) = \mathbb{E}_\mu[\tau] > 0$ because $\tau > 0$ a.s.. Letting Q be the generator of $(V_t)_{t \geq 0}$, we will show below that

$$hQ = 0. \quad (4.6)$$

This identity implies that $h/\mathbb{E}_\mu[\tau]$ is one invariant probability distribution for V . Since π_V is the unique invariant distribution, we deduce that for all $v \in \Omega_V$

$$\frac{h(v)}{\mathbb{E}_\mu[\tau]} = \pi_V(v),$$

which is precisely what we need to prove.

We will derive $hQ = 0$ from the limit

$$\forall v \in \Omega_V : hQ(v) = \lim_{\varepsilon \searrow 0} \frac{[h e^{\varepsilon Q}](v) - h(v)}{\varepsilon}. \quad (4.7)$$

In order to compute the limit we recall $e^{\varepsilon Q}(w, v) = \mathbb{P}_w(X_\varepsilon = v)$ for all $w, v \in \Omega_V$. Therefore

$$[h e^{\varepsilon Q}](v) = \sum_{w \in \Omega_V} h(w) \mathbb{P}_w(V_\varepsilon = v) = \int_0^\infty \sum_{w \in \Omega_V} \mathbb{P}_\mu(V_t = w, \tau > t) \mathbb{P}_w(V_\varepsilon = v) dt.$$

Crucially, the fact that τ is a stopping time for U, V implies that the event $\{V_t = w, \tau > t\}$ is measurable with respect to $(U_s, V_s)_{s \leq t}$. Using that V and U evolve independently and the Markov property for V implies

$$\begin{aligned} \mathbb{P}_\mu(V_t = w, \tau > t) \mathbb{P}_w(V_\varepsilon = v) &= \mathbb{P}_\mu(V_t = w, \tau > t) \mathbb{P}_\mu(V_{t+\varepsilon} = v \mid V_t = w, \tau > t) \\ &= \mathbb{P}_\mu(V_t = w, V_{t+\varepsilon} = v, \tau > t). \end{aligned}$$

Plugging this back in the previous display gives

$$\begin{aligned} [h e^{\varepsilon Q}](v) &= \int_0^\infty \mathbb{P}_\mu(V_{t+\varepsilon} = v, \tau > t) dt \\ &= \int_0^\infty \mathbb{P}_\mu(V_{t+\varepsilon} = v, \tau > t + \varepsilon) dt + \int_0^\infty \mathbb{P}_\mu(V_{t+\varepsilon} = v, t \leq \tau \leq t + \varepsilon) dt \\ &=: (I) + (II). \end{aligned} \quad (4.8)$$

The first term is

$$(I) = \int_\varepsilon^\infty \mathbb{P}_\mu(V_t = v, \tau > t) dt = h(v) - \int_0^\varepsilon \mathbb{P}_\mu(V_t = v, \tau > t) dt,$$

so

$$\frac{(I) - h(v)}{\varepsilon} \rightarrow -\mathbb{P}_\mu(V_0 = v, \tau > 0) = -\mathbb{P}_\mu(V_0 = v) \quad (4.9)$$

because $\tau > 0$ always. Regarding the second term, we have

$$(II) = \int_0^\infty \mathbb{P}_\mu(V_t = v, V_\tau = v, \tau \leq t \leq \tau + \varepsilon) dt + \int_0^\infty \mathbb{P}_\mu(V_t = v, V_\tau \neq v, \tau \leq t \leq \tau + \varepsilon) dt. \quad (4.10)$$

For the first term on the right hand side above we obtain

$$\lim_{\varepsilon \rightarrow 0} \frac{\int_0^\infty \mathbb{P}_\mu(V_t = v, V_\tau = v, \tau \leq t \leq \tau + \varepsilon) dt}{\varepsilon} = \mathbb{P}_\mu(V_\tau = v). \quad (4.11)$$

As for the second term in the sum in (4.10) we get

$$\begin{aligned} \int_0^\infty \mathbb{P}_\mu(V_t = v, V_\tau \neq v, \tau \leq t \leq \tau + \varepsilon) dt &= \int_0^\infty \mathbb{E}_\mu[\mathbb{P}_\mu(V_t = v, V_\tau \neq v, \tau \leq t \leq \tau + \varepsilon \mid \tau)] dt \\ &= \int_0^\infty \mathbb{E}_\mu[\mathbf{1}(\tau \leq t \leq \tau + \varepsilon) \mathbb{P}_\mu(V_t = v, V_\tau \neq v \mid \tau)] dt. \end{aligned}$$

On the event $\{\tau \leq t \leq \tau + \varepsilon\}$ in order to have $V_t = v$ and $V_\tau \neq v$, there must exist at least one jump of the Markov chain in the time interval $[\tau, t]$, which on this event has length less than ε . Therefore, we obtain that on the event $\{\tau \leq t \leq \tau + \varepsilon\}$

$$\mathbb{P}_\mu(V_t = v, V_\tau \neq v \mid \tau) = O(\varepsilon).$$

Therefore we deduce

$$\int_0^\infty \mathbb{E}_\mu[\mathbf{1}(\tau \leq t \leq \tau + \varepsilon) \mathbb{P}_\mu(V_t = v, V_\tau \neq v \mid \tau)] dt = O(\varepsilon^2).$$

Hence this together with (4.11) gives that

$$\frac{(II)}{\varepsilon} \rightarrow \mathbb{P}_\mu(V_\tau = v) \text{ as } \varepsilon \searrow 0.$$

Combining this with (4.9) and (4.8) gives:

$$\frac{[h e^{\varepsilon Q}](v) - h(v)}{\varepsilon} \rightarrow \mathbb{P}_\mu(V_\tau = v) - \mathbb{P}_\mu(V_0 = v)$$

Our assumption that $\mathbb{P}_\mu(V_0 = \cdot) = \mathbb{P}_\mu(V_\tau = \cdot)$ implies that the right hand side above is zero. Plugging this back into (4.7) gives $hQ = 0$ and finishes the proof. \square

Proof of Lemma 4.1. We want to eventually use Lemma 4.2 to estimate the integral, which we can rewrite as

$$\int_0^\infty \mathbb{P}_\mu(X_t = Y_t, t < M^{\text{bad}}) dt = \sum_{x \in \Omega} \mathbb{E}_\mu \left[\int_0^{M^{\text{bad}}} \mathbf{1}(X_t = Y_t = x) dt \right]. \quad (4.12)$$

This is a sum of terms of the form demanded by Lemma 4.2: just set $(U_t)_{t \geq 0} = (Z_t)_{t \geq 0}$ and $(V_t)_{t \geq 0} = (X_t, Y_t)_{t \geq 0}$. However, other conditions are needed for this lemma to apply, and one of them clearly fails: the distribution of (X_0, Y_0) is *not the same* as that of $(X_{M^{\text{bad}}}, Y_{M^{\text{bad}}})$. To see this, simply note that whereas $X_0 = Y_0$ under μ (as we will see below), typically $X_{M^{\text{bad}}} \neq Y_{M^{\text{bad}}}$. It turns out that we can circumvent this problem by defining

$$\tau = \inf\{t \geq M^{\text{bad}} : X_t = Y_t\}. \quad (4.13)$$

Note that $X_t \neq Y_t$ for all $M^{\text{bad}} \leq t < \tau$. Therefore

$$\int_0^\infty \mathbb{P}_\mu(X_t = Y_t, t < M^{\text{bad}}) dt = \sum_{x \in \Omega} \mathbb{E}_\mu \left[\int_0^\tau \mathbf{1}(X_t = Y_t = x) dt \right]. \quad (4.14)$$

Clearly τ is a stopping time for $(X_t, Y_t, Z_t)_{t \geq 0}$. We claim that τ and the initial distribution μ satisfy the conditions (1) – (3) of Lemma 4.2.

First, $\mathbb{P}_\mu(\tau > 0) = 1$, since $\mathbb{P}_\mu(X_0 = Y_0 \neq Z_0) = 1$ and $\mathbb{P}_\mu(\tau \geq M^{\text{bad}} > 0) = 1$. Hence condition (1) of Lemma 4.2 is satisfied. Next we show that

$$\mathbb{E}_\mu[\tau] \leq 22t_{\text{hit}}^*, \quad (4.15)$$

which will imply that condition (2) is also satisfied. Recalling that $\tau = \inf\{t \geq M^{\text{bad}} : X_t = Y_t\}$, we have

$$\mathbb{E}_\mu[\tau] \leq \max_{(x,y,z)} \mathbb{E}_{(x,y,z)}[M^{\text{bad}}] + \max_{(x',y',z')} \mathbb{E}_{(x',y',z')}[M^{\text{good}}].$$

The first expectation is at most the meeting time of X and Z , which is independent of Y . Lemma 3.1 implies that, conditionally on Z , this is at most $11t_{\text{hit}}^*$ almost surely, so $\mathbb{E}_{(x,y,z)}[M^{\text{bad}}] \leq 11t_{\text{hit}}^*$. Similarly, M^{good} is the meeting time of X and the independent trajectory Y , and $\mathbb{E}_{(x',y',z')}[M^{\text{good}}] \leq 11t_{\text{hit}}^*$.

Note now that by definition μ is supported on the set $\{(x, x, z) : x \neq z\}$. One can also check that μ is invariant under the action of any automorphism φ of P , i.e. $\mu(x, y, z) = \mu(\varphi(x), \varphi(y), \varphi(z))$. This together with transitivity give that the marginal of μ on the first two coordinates is uniform on the diagonal set $\Delta = \{(x, x) \in \Omega^2 : x \in \Omega\}$.

Similarly the law of (X_τ, Y_τ) is invariant under the action of any automorphism φ . Using transitivity again we obtain that (X_τ, Y_τ) is uniform on Δ . Therefore, condition (3) is also satisfied.

We can now apply Lemma 4.2 to get that

$$\int_0^\infty \mathbb{P}_\mu(X_t = Y_t, t < M^{\text{bad}}) dt = \sum_{x \in \Omega} \frac{\mathbb{E}_\mu[\tau]}{n^2} \leq \frac{22t_{\text{hit}}^*}{n},$$

where the inequality follows from (4.15). This concludes the proof. \square

5 Non transitive chains

The goal of this section is to prove Theorem 1.4. Throughout the section we fix $\varepsilon > 0$ and let $C = 6/\varepsilon$. In what follows K_r is the complete graph on $r \in \mathbb{N} \setminus \{0\}$ vertices.

For $n \in \mathbb{N}$ construct a graph G_n as follows: begin from a clique K_{n+1} and n disjoint copies of K_k with $k = \sqrt{Cn}$. Fix a vertex $v \in K_{n+1}$ and add exactly one edge from v to each copy of K_k . See Figure 1 for a depiction of the graph.

Let Ω be the vertex set of G_n . We call Down the set of vertices belonging to K_{n+1} and Up = $\Omega \setminus \text{Down}$ the rest.

Let P be the transition matrix of a simple random walk over G and π its stationary distribution. Let X, Y and Z be independent random walks starting from π with transition matrix P and speeds $\lambda_X = 1, \lambda_Y = 0$ and $\lambda_Z = 1$.

The idea is that by choosing ε sufficiently small, the stationary measure of Down becomes arbitrarily small. So if we start X, Y and Z according to π , then it is very likely they will all start from different cliques in Up. Let T be the \sqrt{n} -th time that X visits the vertex v . We will show that as $n \rightarrow \infty$ the probability that X and Z collide after time T is arbitrarily small. Moreover, we will show that the probability that X and Y collide before T is arbitrarily small as $n \rightarrow \infty$. Combining these two assertions will complete the proof.

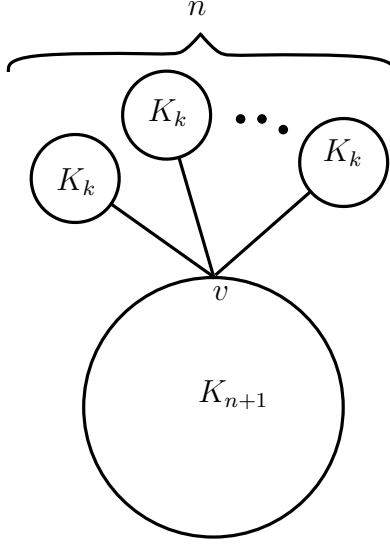


Figure 1: The graph G

For all $r \geq 0$ we define $\tau_v^{(r)}$ to be the time of the r -th visit to v . Formally,

$$\tau_v^{(0)} = \inf\{t \geq 0 : X_t = v\}$$

and for $i \geq 1$ we define

$$\tau_v^{(i)} = \inf\{t > \tau_v^{(i-1)} : X_t = v, X_{t-} \neq v\}.$$

Lemma 5.1. *There exists $\alpha = \alpha(C) > 0$ independent of n such that for all $x, z \in \Omega$ and all $r \geq 1$ we have*

$$\mathbb{P}_{x,z}(M^{X,Z} > \tau_v^{(r)}) \leq (1 - \alpha)^{r-1}.$$

Proof. First note that by the strong Markov property we have for all $r \geq 1$

$$\sup_{x,z} \mathbb{P}_{x,z}(M^{X,Z} > \tau_v^{(r)}) \leq \sup_z \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(r-1)}).$$

Using the strong Markov property again, for $r \geq 1$ we obtain

$$\begin{aligned} \sup_z \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(r)}) &= \sup_z \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(r)} \mid M^{X,Z} > \tau_v^{(1)}) \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(1)}) \\ &\leq \sup_w \mathbb{P}_{v,w}(M^{X,Z} > \tau_v^{(r-1)}) \sup_z \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(1)}). \end{aligned}$$

By induction for all $r \geq 1$ this yields

$$\sup_z \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(r)}) \leq \left(\sup_z \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(1)}) \right)^r.$$

So we complete the proof by showing that

$$\sup_z \mathbb{P}_{v,z}(M^{X,Z} > \tau_v^{(1)}) \leq 1 - \alpha \tag{5.1}$$

for a positive constant α depending only on C .

Let $\tau = \inf\{t \geq 0 : Z_t \in \text{Down} \setminus \{v\}\}$ and fix $w \in \text{Down} \setminus \{v\}$. By symmetry, for all z we then have

$$\mathbb{P}_{v,z}(M^{X,Z} \leq \tau_v^{(1)}) \geq \frac{1}{2} \cdot \min_z \mathbb{P}_{w,z}(\tau \leq \tau_v^{(0)}) \min_{a,b \in \text{Down} \setminus \{v\}} \mathbb{P}_{a,b}(M^{X,Z} \leq \tau_v^{(0)}),$$

where the factor $1/2$ corresponds to the probability that the first time X jumps it goes to $\text{Down} \setminus \{v\}$.

If $X_0 = a \in \text{Down} \setminus \{v\}$, then $\tau_v^{(0)} = \tau_v^X$, and hence if also $b \in \text{Down} \setminus \{v\}$, then

$$\mathbb{P}_{a,b}(M^{X,Z} \leq \tau_v^{(0)}) \geq \mathbb{P}_{a,b}(M^{X,Z} \leq \tau_v^X \wedge \tau_v^Z) = \frac{1}{2}.$$

It remains to show that for a positive constant c_1 we have

$$\min_z \mathbb{P}_{w,z}(\tau \leq \tau_v^{(0)}) \geq c_1 > 0. \quad (5.2)$$

If $z \in \text{Down} \setminus \{v\}$, then this probability is 1 and if $z = v$ it is easily seen to be at least $1/4$. So we assume that $z \in \text{Up}$. Let x be the unique neighbour of v lying in the same clique as z . Then the time τ can be expressed as $\tau = T_{z,x} + T_{x,v} + T_{v,\text{Down} \setminus \{v\}}$, where the time $T_{r,S}$ stands for the first hitting time of S starting from r . Using this, it is then not hard to see that there exists a positive constant c such that uniformly over all $z \in \text{Up}$ we have $\mathbb{E}[\tau] \leq ck^2$. Moreover, if $X_0 \in \text{Down} \setminus \{v\}$, then $\tau_v^{(0)}$ is an exponential random variable with mean n . By Markov's inequality we obtain

$$\mathbb{P}_{w,z}(\tau \leq \tau_v^{(0)}) \geq \mathbb{P}_z(\tau \leq 2\mathbb{E}[\tau]) \cdot \mathbb{P}_w(\tau_v^{(0)} \geq 2\mathbb{E}[\tau]) \geq \frac{1}{2} \cdot \int_{2\mathbb{E}[\tau]}^{\infty} ne^{-ns} ds = \frac{1}{2}e^{-2n\mathbb{E}[\tau]}.$$

Note that this bound does not depend on z . Since $k = \sqrt{Cn}$ and $\mathbb{E}[\tau] \leq ck^2$ the bound in (5.2) follows. \square

Proof of Theorem 1.4. We show that for n sufficiently large, the graph $G = G_n$ satisfies the claim of the theorem.

It is not hard to verify that for n large enough, in G_n we have

$$\pi(\text{Down}) \leq \frac{2}{C}.$$

Let A be the set of pairs (x, y) such that $y \in \text{Up}$ and x is not in the same clique as y . Then let $E = \{(X_0, Y_0) \in A\}$. By the preceding bound $\mathbb{P}(E^c) \leq 3/C = \varepsilon/2$ for n large enough. We then have

$$\begin{aligned} \mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) &\leq \mathbb{P}(M^{X,Y} \leq M^{X,Z}) = \mathbb{P}(M^{X,Y} \leq M^{X,Z}, E) + \mathbb{P}(M^{X,Y} \leq M^{X,Z}, E^c) \\ &\leq \sup_{x,y,z: (x,y) \in A} \mathbb{P}_{x,y,z}(M^{X,Y} \leq M^{X,Z}) + \frac{\varepsilon}{2}. \end{aligned} \quad (5.3)$$

Therefore, it suffices to upper bound the last probability appearing above. Fix any $z \in \Omega$ and $(x, y) \in A$. For r to be determined later we have

$$\mathbb{P}_{x,y,z}(M^{X,Y} \leq M^{X,Z}) \leq \mathbb{P}_{x,y}(M^{X,Y} \leq \tau_v^{(r)}) + \mathbb{P}_{x,z}(M^{X,Z} > \tau_v^{(r)}). \quad (5.4)$$

Because Y is not moving, we have $M^{X,Y} = \tau_y$, where $y = Y_0$. Since x, y are not in the same clique of Up , if $\tau_y \leq \tau_v^{(r)}$, then there exists $1 \leq i \leq r$ such that $\tau_v^{(i-1)} < \tau_y \leq \tau_v^{(i)}$. By the strong Markov property and union bound we obtain

$$\mathbb{P}_{x,y}(M^{X,Y} \leq \tau_v^{(r)}) \leq r\mathbb{P}_v(\tau_y \leq \tau_v^{(1)}) \leq \frac{r}{2n},$$

since, when $X_0 = v$, in order to hit $y \in \text{Up}$ before returning to v , the first time X moves it must jump into the clique that contains y .

Using the above bound and Lemma 5.1 in (5.4) we deduce

$$\mathbb{P}_{x,y,z}(M^{X,Y} \leq M^{X,Z}) \leq \frac{r}{2n} + (1 - \alpha)^{r-1}.$$

Taking $r = \sqrt{n}$ or any other function of n that goes to infinity slower than n gives that

$$\mathbb{P}_{x,y,z}(M^{X,Y} \leq M^{X,Z}) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

We conclude from (5.3) that

$$\mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) < \varepsilon$$

and this finishes the proof. \square

6 Sharpness of Conjecture 1.5

In this section we describe the example pointed out by Alexander Holroyd, mentioned in the Introduction, of a family of transitive graphs for which $\mathbb{P}(M^{\text{good}} \leq M^{\text{bad}}) \leq 1/3 + \delta$. In what follows we take $\lambda_X = \lambda_Y = 1$ and $\lambda_Z = 0$.

To construct the example, fix $\varepsilon \in (0, 1)$ and consider the chain with state space $\{0, 1\}^n$ in which the j 'th coordinate changes value (from 0 to 1 or vice-versa) at rate $q_j = \varepsilon^{j-1}(1 - \varepsilon)/(1 - \varepsilon^n)$; note that $\sum_{i=1}^n q_i = 1$. The idea is that for small ε , earlier coordinates change state much more quickly than later coordinates, so the primary obstacle to both meeting and hitting is simply the largest coordinate in which the value differs. For $u, v \in \{0, 1\}^n$, let $k(u, v) = \max\{i : u_i \neq v_i\}$, or $k(u, v) = 0$ if $u = v$.

We claim that for $x, y, z \in \{0, 1\}^n$, if $k(x, y) > \min(k(x, z), k(y, z))$ then $\mathbb{P}_{x,y,z}(M^{\text{good}} < M^{\text{bad}}) < 2\varepsilon n$. Assuming this, and taking $\varepsilon = \delta/(2n)$, it follows by symmetry that, starting from stationarity,

$$\mathbb{P}(M^{\text{good}} < M^{\text{bad}}) \leq \delta + \mathbb{P}(k(X_0, Y_0) < \min(k(X_0, Z_0), k(Y_0, Z_0))) < \delta + \frac{1}{3}.$$

It thus remains to prove the preceding claim.

Fix $x, y, z \in \{0, 1\}^n$ with $k(x, y) > \min(k(x, z), k(y, z))$, and assume by symmetry that $k(x, z) < k(x, y)$. For $1 \leq k \leq n$, let $\tau_k = \min\{t : X_t^{(i)} = z_i, 1 \leq i \leq k\}$ be the first time that X_t and z agree in the first k coordinates. It is convenient to set $\tau_0 = 0$. Also, let $\sigma_k^X = \min\{t : \exists i \geq k, X_t^{(i)} \neq X_0^{(i)}\}$ be the first time one of the last $n - k + 1$ coordinates of X changes, and define σ_k^Y accordingly.

We will show that for all $1 \leq k < n$,

$$\mathbb{P}(\tau_k < \sigma_{k+1}^X) \geq (1 - \varepsilon)^k \geq 1 - k\varepsilon. \quad (6.1)$$

Note that τ_k , σ_{k+1}^X and σ_{k+1}^Y are all independent. Furthermore, σ_{k+1}^X and σ_{k+1}^Y are identically distributed, so if the preceding inequality holds as written then it also holds with σ_{k+1}^Y in place of σ_{k+1}^X . We finish proving the claim assuming that (6.1) holds, then conclude by proving (6.1).

At time $\tau_{k(x,z)}$, the first $k(x, z)$ coordinates of X agree with those of z . If $\tau_{k(x,z)} < \sigma_{k(x,z)+1}^X$ then the remaining coordinates of X and z also agree (because they did at time 0 and they have not changed),

so $M^{X,Z} = \tau_{k(x,z)}$. Similarly, if $\tau_{k(x,z)} < \sigma_{k(x,z)+1}^X$ and $\tau_{k(x,z)} < \sigma_{k(x,z)+1}^Y$ then $M^{\text{bad}} < M^{\text{good}}$. It then follows, using (6.1) and the subsequent observation, that

$$\mathbb{P}_{x,y,z}(M^{\text{good}} < M^{\text{bad}}) \leq \mathbb{P}(\sigma_{k(x,z)+1}^X \leq \tau_{k(x,z)}) + \mathbb{P}(\sigma_{k(x,z)+1}^Y \leq \tau_{k(x,z)}) \leq 2k(x,z)\varepsilon < 2n\varepsilon,$$

as claimed. It thus remains to prove (6.1). In what follows we write $\sigma_k = \sigma_k^X$.

Fix $1 \leq k < n$, and note that σ_k is exponential with rate $\sum_{j=k}^n q_j = \varepsilon^{k-1}(1 - \varepsilon^{n+1-k})/(1 - \varepsilon^n)$. Furthermore, $\sigma_k < \sigma_{k+1}$ precisely if the k 'th coordinate of X changes before any larger coordinate. It follows that $\mathbb{P}(\sigma_k < \sigma_{k+1}) = q_k / \sum_{j=k}^n q_j$.

Suppose that $X_0^{(k)} = z_k$. In this case to have $\tau_k < \sigma_{k+1}$ it suffices that $\tau_{k-1} < \sigma_k$, so

$$\mathbb{P}(\tau_k < \sigma_{k+1} \mid X_0^{(k)} = z_k) \geq \mathbb{P}(\tau_{k-1} < \sigma_k).$$

If $X_0^{(k)} \neq z_k$ then the k 'th coordinate must change before time τ_k , so to have $\tau_k < \sigma_{k+1}$ it is necessary that $\sigma_k < \sigma_{k+1}$. By the strong Markov property, we then have

$$\begin{aligned} \mathbb{P}(\tau_k < \sigma_{k+1} \mid X_0^{(k)} \neq z_k) &= \mathbb{P}(\sigma_k < \sigma_{k+1}) \mathbb{P}(\tau_k < \sigma_{k+1} \mid X_0^{(k)} = z_k) \\ &= \frac{q_k}{\sum_{j=k}^n q_j} \cdot \mathbb{P}(\tau_k < \sigma_{k+1} \mid X_0^{(k)} = z_k) \\ &\geq \frac{1 - \varepsilon}{1 - \varepsilon^{n+1-k}} \mathbb{P}(\tau_{k-1} < \sigma_k). \end{aligned}$$

We thus have the unconditional bound

$$\mathbb{P}(\tau_k < \sigma_{k+1}) \geq \frac{1 - \varepsilon}{1 - \varepsilon^{n+1-k}} \mathbb{P}(\tau_{k-1} < \sigma_k) > (1 - \varepsilon) \mathbb{P}(\tau_{k-1} < \sigma_k),$$

This bound holds for all $1 \leq k < n$; since $\tau_0 = 0$ we also have $\mathbb{P}(\tau_0 < \sigma_1) = 1$, and (6.1) follows.

Acknowledgements

The authors thank the Bellairs Institute, Microsoft Research and the Isaac Newton Institute, where parts of this work were completed.

References

- [1] David Aldous and James Fill. *Reversible Markov Chains and Random Walks on Graphs*. In preparation, <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [2] David J. Aldous and Mark Brown. Inequalities for rare events in time-reversible Markov chains. I. In *Stochastic inequalities (Seattle, WA, 1991)*, volume 22 of *IMS Lecture Notes Monogr. Ser.*, pages 1–16. Inst. Math. Statist., Hayward, CA, 1992.
- [3] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

- [4] Roberto Imbuzeiro Oliveira. On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.*, 364(4):2109–2128, 2012.
- [5] Perla Sousi and Peter Winkler. Mixing times and moving targets. *Combin. Probab. Comput.*, 23(3):460–476, 2014.